

Additional File 3
Fasta sequences of retrieved platyhelminth Dnmt and MBD homologs

NCBI database

Neobenedenia melleni Dnmt (NmDnmt; GW920589.1)

```
>Nm_Dnmt
MKTMEINVAELSGIGGMRLALKKSTLQFNVLQAYDIHDGANDIYNYNKEKCAKNRDIK
ALHCEESVLRCELTWTMSPPCQPYTRLGNMKQSKDTRSEPFAHIELIRKMSPKLIFLEN
VQGFEESDSHEVLISCLIMSYDIKQYLLTPLQFGVPNCRTRFYLIAKLSPETPSLKIKT
TLDNNLPRLPGCCCKYCSTEYNESGEEITNYLECNENILNFVKDEDLYRYFRAMDVVSPKD
SKTKCFTKGYSKRFTGTGSFLQCCENPIKHDDIDEKRNSF
```

Paragonimus westermani MBD (PwMBD; AT007407)

```
>Pw_MBD
MLNHQSNQPKRNVYINVPISTAVAPPCTSNGHSQVIPISLPAWRREECTRPSGLGTGK
TDVFYMSPSGQKFRTKQEMKAVLGDACDLTLFDWRSGKFLAPSQVRRPDDGSDPTASKV
PRVDGCSVLTRRTNPETFVPIIRSHPECKRTDVRNAQHETPHQLFWEKRTGKHAVDP
DTGEALKPLNLPNTIQTAGVPGYQPAQMVSQSLVNAIAS
```

Schistosoma japonicum Dnmt (SjDnmt; FN238874 & FN328873)

```
>Sj_Dnmt
MRVLELYSGIGGMHAFQMSLKHDEVVAIEINDIATDVYKHNFPNTLTLNRVIESFSAN
SLCSFNSNVWSLCPPCQPFTRLGKRMCEVDNRSSFFHVLDLISILKPEGIILENVKGFE
HSEPWRRLMKVLNACDYEYRQFLLSPLOFGIPNCRLRFYLVARSSSSWNSNLKMGKSES
IDLRRPDISPVLPGCRCASCCGLVSHIEHTDDKTDYIQFCLPVSFLLTANDREELNFL
DRKCLERYFRVLDIVRSCDKKTRCFKGYSKRLEGTSVIQTSMEDETGDKITSYYESNK
ENKQAVLQYAEQLKLRFHSREVANMLCFPKRDFPERVTEKQRLLLGNSVNVLVSHL
I
```

Schistosoma japonicum MBD (SjMBD; AAW26585.1)

```
>Sj_MBD
MMSVTGDNKLCIMQNQSIPANQLNNSSQTKRSSYSYNKLQTTSSSPQQQQQTWPSSLPP
GWKREECMRPNGLGTGKSDVYYISPQNQVVRNKQEMQTIQDGYDIAFDWRIGKFNMNN
NNKSKRLEEVTADTAAKVPRLDNQYNLPVRRCLFSSTTNEIKPVIVRSHPECKRTDVK
NTNHEIPHQLFWEKRLADHMAINPDTGEFPFKPISLPKGMCRRILAGVPGYQPTQLIQSV
SHALSSSSNNNNSPIIGQEQQPSAIEKNPCVIINHLQPMIKTFIVTDDDIRQEARVKE
LRKKLETARKKLHPRYETEREQ
```

WTSI GeneDB (<http://www.genedb.org/>)

Echinococcus multilocularis Dnmt (EmDnmt; EmuJ_001185500)

```
>Em_Dnmt
MRVLELYSGIGGMRAALELSGLAFEIARALDVNDLANSVYNSAYNCSDASNRALESSEE
ECLLFGADLWTMSPPCQPFTRMGNQKRGKDTRCSSLLIIQLIRRIKPPAVMLENVKGF
GSDAWCAILEALIACDYDVRFLTPLQFGIPNCRLRYYLVARLRTHNKQRMFSFGTSVDD
LKGSFEKLASIIRTPPCDVSPMPNCECGVCTNKVGIASTENHFLEYIPFCHPIADYLL
PECEIPEDLFLGQNELEKYYTILDIVTPENRKSACFTKGYAQRFEGTGSFLEVQPSRSN
STSFPKIRAFHSKEIARLLCFPEYFDFPATVSEKQRRLLGNSLNVLVVAHIMNWAFSTV
```

Echinococcus multilocularis MBD (EmMBD; EmuJ_001033700)

```
>Em_MBD
MATHPKYVSQSPLSVNKQIAQNMMFQKVTHNTPITQKSASALPVWKREEILRPDGLS
SGKTVVFYVSPNGMKIRTQKELQNALGDKFDISSLFDWRTGKFSSLLKHRIPNSGDQNSV
KSSFFRSITSSECKTEDIEYDWFQRNTTPNFPPEPVIVLNHNDNSRSDKSPVKESLKQLFS
ERRLSDRYAVDHDRGETLIRSTLPPGIESAGVPGYNKSQLLQNLIVSLSKHPITGQER
TVEQNPCAMLNQNQPYVKNFVITDEDIRKQYQRVKDIRDRLQAARKVYPEFLFDLSLSRW
LQTAIKPSVSSFTIDASGTPSTLYIQFVHIVVGNRFVSRH
```

Echinococcus granulosus Dnmt (EgDnmt; EgrG_001185500)

```
>EgDnmt
MRVLELYSGIGGMRAALELSGLPFEITRALDVNDLANSVNSAHNCSDASNRALESSEE
```

ECLLFGADLWTMSPCQPFTRMGNQKRGKDTRCSSLLIIQLIRRIKPPAVILENVKGFE
GSDAWRIILEALIACDYDVRFLTPLQFGIPNCRLRYYLVARLRTHSKQRMFSFGTSVDD
LKGSFEELASAVIRTTPCDVSPPNCECVCTNKVGGIASTENHFLEYIPFCRPIADYLL
PECEIPEDLFLGQNELEKYYTILDIVTPENRKSACFTKGYAQRFEGTGSFLEVQPPLHSN
STSFPKIRAFHSKEIARLLCFPEYFDFPATVSEKQRRRLGNSNLVVAHIMNWAFSTV

Echinococcus granulosus MBD (EgMBD; EgrG_001033700)

>EgMBD

MATHPKYVSQSPLSVNKQIAQNMMFQKVTTINHAVPITQKSASALPVWKREEILRPDGLN
SGKTVVVIYVSPNGMKIQTQKELQNALGDKFDISSLFDWRTGKFSSLLKHRIPNSGDQNSV
KSSFFRSITSSECKTEDVEYDWFQRNTTPNFPPEPVIVLNHNDNSRSDKSPVKESLKQLF
ERRLSDRYAVDHRDGKTLIRSTLPPGIESAGVPGYNKSQQLQNLIVSLSSKGHPITGQER
TVEQNPCAMLNQNQPYVKNFVITDEDIRKQYQRVKDIRDRLQAARKVCPEFFFDSLRSRW
LQTAIKPSVSNFTIDASDTPSTLYIQFVHVIVGNRFVCVRRH

Hymenolepis microstoma Dnmt (HmDnmt; HmN_000063350)

>HmDnmt

MRTLEFYSGIGGMHAALKTSGLPHEVIKAFFEINDSANSIYNEVDCSYATNRITIESLSEM
ECLAYDAELWTMSPCQPFTRLGNKKQGLDSRSSSLYNNIIDLIKLIKPPVIFLENVKGF
GSDIWKKLIETLILCNYDIRQFLLTPLOFGVPNCRLRYYLVAKLRTNDDQTMFKFGE
HLQSFLENIEDNIIRLPIFDVEPMPGCDCCVCTKQVKHIDSIEENDYEQYLPFCSPVSSY
LPDNELPEDIYLGQDILKKYYRILDIIEPESHKTACFTKGYGRRYEGTGSFLSLSKVGDI
GEHHFPQMRRFHSKEVAKIMCPDNFEFPNFTEKQRLKLLGNSINVRVVAHIMHTLSP

Hymenolepis microstoma MBD (HmMBD; HmN_000163900)

>HmMBD

MAQFKSPSTPVPRPKMEAQQILINRLKTIKLNFITMIQKTGFNHQSSAIQKTLVPLSG
WRREEVIRPSGLNTGKNLVFYVSPDGKTYKTLKELQLGLGSKYDLSTFDWRTGKFLLTSL
KYKPGNEDHVNHKRSKPYDMNDWFQRGSAIKFPKIELKSNHGLSKRIDKPEAQECLRQ
LFAERRLFDRCPVDHRDRKTIEATLPPGVESCGVPGYNKMQLMQNLIVSLYTKSGPISG
QERCVEQNPCAMINQNQPFVKSFIVIDEDEVRKQQHRIAEEQEF

The Taenia solium Genome Project (<http://www.taeniasolium.unam.mx/taenia/>)

Taenia solium Dnmt (TsDnmt; TsM_000460600 with added/altered exons from genomic scaffold TsM1_s601)

>TsDnmt

MRILELYSGIGGMRAALESLGLPKIICKAFDINDLANFVYNSVYKCSASNRALESILSEA
ECLLYDADLWTMSPCQPFTRMGNQKRGQDVRCSSLVVIHLIRRTKPPAILLENVKGF
GSDAWRYILDALIECDYDVRQFLLTPLOFGIPNGRLRYYLVARLRTRSEQQMFSFGDSVD
LNGCLREPTSAIMRTPPCGVSPMPNCECPVCTNKIRQVLLRYMRFVFSFSDMVSPENHFL
EYLPFCRPIADFLTDLEIPESLFLGQNVLKEYYTLDIVTPESRKSACFTKGYAHRFEG
TGSFLEMSQSSHSGLTPFPKIRAFHSKEIARLLCFPEHDFPEEVSEKQRRRLGNSINV
LVVAHVMNWTLSV

Taenia solium MBD (TsMBD; TsM_000235700 with altered exon based on cestode MBD gene structure)

>TsMBD

MPITQKSVSALPVWKREEILRPDGLNSGKTIVFYVSPGMKVRTQKELQNVLGDKFDIS
SFDWRTGKFSSPLLKHRVPNLGEQSSVKCKTEDAECDWFQRNTTPNFPKPVIVLNHTDS
NRSDKSPIESLKQLFSSRLSDRYAVDHRDGKTLMRSTLPPGIEAAGVPGYNKSQQLQ
LIVSLSSKGHPITGQERTVEQNPCAMLNQNQPYVKNFVITDEDIRKQYQRVKDIRDRLQA
ARKVTSSGHQP

SchistoDB beta version (<http://beta.schistodb.net/schistobeta/>)

Schistosoma haematobium Dnmt (ShDnmt; Sha_108347 - added exon positions from Scaffold471 319482 – 319631, 319663-319768bp)

>Sh_Dnmt

MRVLELYAGIGGMHIAFKGSTVKHEVVAAVEINDVATDVYKYNFPNTLTLNRVIESFSPDY
VCSLNANIWSLCPPCQPFTRLGKRMCEADKRSSSSFFHVLDDISILKPTGIILENVKGFEH
SEPWROIIEVLSNCDYEYRQFLSPLQFGIPNCRLRFYLLARLRSNSNFKMGQSES
DMRPPVDAAPMLPGCQCTSCSGVISHIEHTDDNFTYEIQFCRPISEFVLVPSDSPKELYFL
DEKCLQRYFRVLDIVRSCDKKTRCFTKGYSKRLEGTSVQFTSMENEICVNYRLNHKQRV
DKFSLEIFQSMLTSEKITNYYEANKEDEQAVLQYAKLLKLRFHSREVANMMCFPKSFDF

PEHITEKQRLLLGNSVNILVVSHLIYWVFGT

SmedGD v1.3.1.4 (<http://smedgd.neuro.utah.edu/>)

Schmidtea mediterranea Dnmt (SmdDnmt; mk4.000626.03.01)

>Smd_Dnmt

```
MKVIEFFSGIGGIKCALNGNILSNAISLTDFINDLANSVYRYNFTSPCYNKVIESTSAEY
LISLNADMWTMSPPCQPYTRNGNMMDDPRTAAMKHTLYLISQVRPHYIFFENVKGFE
SNGRKMLVISLSESAYSFQSQEFLSPLQFGVPNSRLRYYLIAKLEGKGNLMDLEAISY
KPYFDRKLKYQCNCPCVCSGRSRSLENDHVNHFERNLEFCDRISAYLEADNLAEPEGHKLI
DEKVLEKGFSKLDIVTESSNKTCCEFCKYAKKIEGSGSYYQMTSCEEAHQLKQLLLNGDI
SSLDYAKRLQLRYFSPREIANFMCFPQSFRFPETVTRAQRYRLLGNSVNVKVVAHVLHWL
ISA
```

Schmidtea mediterranea MBD (SmdMBD; mk4.000465.10.01)

>Smd_MBD

```
MSKQSFPSNTNLNSKKANSATQPTRTSVFNATAVNHQTTSLPPGWKREEILKPTGLYAG
KTDIYYTSPGQKIRSKAEMQKILGGSYDVSNFDWHSGKFVNSRKRPSTNENDSTSCKLM
KMIEEYQQCLVRSGPSLFCKDKLPPIETTRPDSRINNDAVGGNDNQEVPKQIFWHKRLAD
NYAFNPNTNEKVRSELVPKEFQNAGFPGIDNNQLFQSIVSCIYNSTKNNIPVLGQSNHKA
SIEKNPCLLITANQPLVKQVVITDDIINRQEAKVKELRRKLELVRRTLLPKLSVDEAA
```

Gasser Laboratory database (<http://www.gasserlab.org/>) & NCBI sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>)

Clonorchis sinensis Dnmt (CsDnmt; Contig6449)

>Cs_Dnmt

```
MRVLELYSGIGGMHCAFKA
QLDFQVVAIAIDINDTANRVYRANFAHTYAPNRVLESLSIEEVTAFNADMWSMSPPCQPFT
RLGNQKHEEDNRSASFFYVLGLIAIRPKFILLEENVKGFEHTEPWCQFLKVLQTCGYRYQ
QFLLTPLQFGVPNCRLRYYLVASSSSDSACSGLFASHVHDSDHANTIHLLIPPADLPLPG
GECAVCLGHVSHITKPDENFDDYLPYCRISPIDSYLAQQQQHSKLDFLDHNCCLKRYFHVDI
VRPCDRKSRCFTKGYQKRIEGTGSVLQTATDTLTSEEINAQWITARSDPETLMSLAKRLQ
LRFFHSREVANLLCFPQTYNFPEDVTEKQIRLLGNSVNVLVVAHLIHWAFCGDLHX
```

Clonorchis sinensis MBD (CsMBD; Contig1306)

>Cs_MBD

```
MLNSQPSGTHSSIGQQKRNTYTHLSKGQSQPHSIVPGGHSVPTFIGLPSGWKREECTRPC
GLSTGKTDVYYISPNGQKVRTKHEMKFLDDAYDISTFDWRTGKFQMSASKVRRTDDSS
PTSAKVPRMDSAFVRRTPPDVFTPIVVRSHPECKRTDXQEPPGASSSAXFGKNVLKTL
MQSIQKQVRLFKQLSLPKDIQSAGVPGYEAPQLVQSLVNALATKSSPITGQEQPVSafeK
NPCVAANNLQPMIKTYIVSDEDVRQEARVRELRRKLELARKKLNPRYSSERHXVEQLDA
EWPERVIALPNSHCTYACLSVPVYNAYLVALIVLIVQLPHVYTMQNCNLANLHSGYFLGQM
HELCCFEQLQSDVGVYPWL
```

Opisthorchis viverrini Dnmt (OvDnmt; Contig11296)

>Ov_Dnmt

```
MRVLELYSGIGGMHCAFKMSQLDFQV
VAIIDINDTANRVYRANFSHTHAPNRVLESLSIEEVTSFTADMWSMSPPCQPFTRLGNQK
YEEDNRSTSFFYVLGLIAIRPKFILLEENVKGFEHTEPWRQFLKVLHNCGYRYQQFLLTP
LQFGVPNCRLRYYLVASSSSDSASSGLFAPHVHDSSHANTIHIVPPADLPLPGCECPVC
LGHVNHITKPDENFDEYLPYCRISPIDSYLTQHQHQHSNLDFLDDSCLRRYFHVLDIRPCDH
KSRCFTKGYQKRIEGTGSVLQTATDPLTGEEIKRPMNDCPSPETQMSLAKRLQLRFFHS
REVA
```

Opisthorchis viverrini MBD (OvMBD; Contig280)

>Ov_MBD

```
MLNSQPSGTHSNIGQQKRNTYTHLSKGQSQPHSILPGGHSVPTFIGLPSGWKREECTRPC
GLSTGKTDVYYISPNGQKVRTKHEMKFLDDAYDISTFDWRTGKFQMSASKVRRTDDSS
PTSAKVPRMDSAFVRRTPPDVFTPIVVRSHPECKRTDVKNLNQEPPHQLFWEKRFKGT
AVDPETGEAFKQLSLPKDIQSAGVPGYEAPQLVQSLVNALATKSSPITGQEQPVSafeKN
PCVAANNLQPMIKTYIVSDEDVRQEARVRELRRKLELARKKLNPIRVNDTDRIGMT
```

Fasciola hepatica genome database (<http://worm1.liv.ac.uk/>.)

Fasciola hepatica Dnmt (FhDnmt; genome sequence (scaffold15595) used with one Gasser Laboratory database transcript transcript (Fh_F hep_A c30799) used to confirm.

```
>Fh_Dnmt
MRVLELYSGIGGIHCALKKSCLSYSIRIAIDVNDLASSVYRHNFPQSPASNRVIEGLSIEEVSRMSADLWTMSPPC
QPFTRLGNRRDEADNRSVSFFHVLTIAIRPSFIFLENVRGFESSESCRELLKTLDSNEYLYQQFLTPLQFGIP
NCRLRFYLVAWRGPTCTNPFKTDDSSLICILDTPPIHLPPLPDCVCSCNGCVRVTANNDNFSEYLPYCQPVANFV
HSLNEHEEKSCSITSCLKRYFYVMDIVRPCDRKTRCFTKGYTKRIEGTGSVLQTSRDLTSIDLHISNYQSVEHD
DALVSFTRTLGLRFFHSREVANLMCFPSDYEPENVTEKQRIRLLGNSVNHHVAHLMFWAFGDRNFPIKPT
```

Fasciola hepatica MBD (FhMBD; genome sequence (scaffold35919) used with Gasser Laboratory database transcripts (FhContig4567 and FhContig3923 used to confirm.

```
>Fh_MBD
MQPQNVSHPNLMQPKRNTYLTFPKTQSFSAMSSNGQQVMQITLPSGWKREALRPTGLGTGKTDVYYISPQGQ
KVRTKQEMKMLLGDKYDIHFFDWRSKGFLSQPPKIPTLQKIVQKYHQHQKIPHLDTSPNPLQRRTSPPDTIKPVVIR
SHPNCKRADMRNLMSMEEPRQLFWEKRLADQVAIDSETNEPCKPLSLPRGVQSAVGPGYRSQQLINSLLYAIATKTS
PIAGQEQAQPSAIEKNPCVAVNTLQPMIKTYIVSEEDIRRQEMRVKELRRKLEMARKRLHPRYATEREG
```

***Macrostomum lignano* genome initiative (<http://www.macgenome.org/>)**

Macrostomum lignano Dnmt (M1Dnmt: RNA815_29248 and RNA815_19240 with the genome sequences (scf25220075753149 & deg 2520075300551) used to confirm

```
>M1Dnmt
MHQGGGEFRVAEFFSGIGGMHCALQLRQLVGEREFSYRIAGAFDVNNIANSVYERNFPGVRCVTTSLDSIGLARI
RALAANLWLMSPPCPQFSRLGNRRGLDDVRCRPLLRLLEVLASLPDNERPSRLLLENVQGFEASPAERLLECLNK
CGYEYRQLLLSPLQFGVPNSRLRFYLLAWRPGAPEPAAARLLPANCIDRPPSAATSQPLLPSCCCPVCSGRLQPG
QLTHDSHLCYRPFARPLSDYLSGPADAEPELYLTSRQLEQFFPVLDIRQPRSATCCFTKGYKRVEGTGSVLT
RRLAFGKPSHDRVKAGACPPRTHGEVSHADAAAIDASCSSLAPDDLSDDDINNGEFVDSAGGSGDAETMQAYAE
SLGLRYFSAAEVAGLMFPPGYSFSTSLLCSRQAYQLLGNSVNVLVIAHLLSYLLS
```

Macrostomum lignano MBD (M1MBD; RNA815_4367 and RNA815_18637 with the genome sequences (scf250075747452) used to confirm

```
>M1MBD
WIEKVFFLCLEQKIQRSTKPVSMATNLTSVVG LLVMPPGWKREESLKQSGLSSGKTEVYYVSPGGRKLNRVELQRALG
DRYDLSNFDWRTGRFHPYARPQQHQHQQASLQSAGRIKTPADMSSAAAAAA
```